Loop-loop interactions govern multiple steps in indole-3-glycerol phosphate synthase catalysis

Margot J. Zaccardi, Kathleen F. O'Rourke, Eric M. Yezdimer, Laura J. Loggia, Svenja Woldt, and David D. Boehr*

Department of Chemistry, The Pennsylvania State University, University Park, Pennsylvania 16802

Received 8 November 2013; Revised 22 December 2013; Accepted 29 December 2013

DOI: 10.1002/pro.2416

Published online 8 January 2014 proteinscience.org

Abstract: Substrate binding, product release, and likely chemical catalysis in the tryptophan biosynthetic enzyme indole-3-glycerol phosphate synthase (IGPS) are dependent on the structural dynamics of the $\beta 1\alpha 1$ active-site loop. Statistical coupling analysis and molecular dynamic simulations had previously indicated that covarying residues in the $\beta 1\alpha 1$ and $\beta 2\alpha 2$ loops, corresponding to Arg54 and Asn90, respectively, in the *Sulfolobus sulfataricus* enzyme (sslIGPS), are likely important for coordinating functional motions of these loops. To test this hypothesis, we characterized site mutants at these positions for changes in catalytic function, protein stability and structural dynamics for the thermophilic sslIGPS enzyme. Although there were only modest changes in the overall steady-state kinetic parameters, solvent viscosity and solvent deuterium kinetic isotope effects indicated that these amino acid substitutions change the identity of the rate-determining step across multiple temperatures. Surprisingly, the N90A substitution had a dramatic effect on the general acid/base catalysis of the dehydration step, as indicated by the loss of the descending limb in the pH rate profile, which we had previously assigned to Lys53 on the $\beta 1\alpha 1$ loop. These changes in enzyme function are accompanied with a quenching of ps-ns and μ s-ms timescale motions in the $\beta 1\alpha 1$ loop as measured by nuclear magnetic resonance studies. Altogether, our studies provide structural, dynamic and functional rationales for the coevolution of residues on the $\beta 1\alpha 1$ and $\beta 2\alpha 2$ loops, and highlight the multiple roles that the $\beta 1\alpha 1$ loop plays in IGPS catalysis. Thus, substitution of covarying residues in the active-site $\beta 1\alpha 1$ and $\beta 2\alpha 2$ loops of indole-3-glycerol phosphate synthase results in functional, structural, and dynamic changes, highlighting the multiple roles that the $\beta 1\alpha 1$ loop plays in enzyme catalysis and the importance of regulating the structural dynamics of this loop through noncovalent interactions with nearby structural elements.

Keywords: enzyme mechanisms; enzyme kinetics; nuclear magnetic resonance; protein dynamics; statistical coupling analysis; amino acid networks; metabolism; protein engineering

Additional Supporting Information may be found in the online version of this article.

Grant sponsor: NSF; Grant number: MCB1053993.

*Correspondence to: David D. Boehr, Department of Chemistry, The Pennsylvania State University, 107 Chemistry Building, University Park, PA. E-mail: ddb12@psu.edu

Introduction

The tryptophan biosynthetic enzyme, indole-3-glycerol phosphate synthase (IGPS), catalyzes the ring closure of 1-(*o*-carboxyphenylamino)-1-deoxyribulose 5-phosphate (CdRP) to indole-3-glycerol phosphate (IGP). IGPS is a validated antibiotic target, including for *Mycobacterium tuberculosis*,^{1–3} and has been frequently used as a starting protein framework

Figure 1. Catalytically important amino acid residues involved in the dehydration step of IGPS are found on the $\beta 1\alpha 1$ and $\beta 2\alpha 2$ loops, which interact through a hydrogen bond between Arg54 and Asn90. (A) In the chemical mechanism for ssIGPS, proton transfer from Lys110 initiates ring closure and decarboxylation. The anthranilate moiety is then transferred to a second site where Glu51 and Lys53 act as the active-site base and acid, respectively. The role of the $\beta 1\alpha 1$ and $\beta 2\alpha 2$ loops including the interaction between Arg54 and Asn90 is examined herein. (B) Close-up of the active-site of ssIGPS indicating important amino acid residues and the $\beta 1\alpha 1$ and $\beta 2\alpha 2$ loops. (C) The interaction between the $\beta 1\alpha 1$ and $\beta 2\alpha 2$ loops is thought to have functional significance in IGPS since it is coevolving amongst IGPS enzymes and exhibits correlated motion in MD simulations. This interaction is in close proximity to the conserved residues Lys53 and Phe89.

in the engineering of novel enzyme activities.^{4,5} We have suggested that IGPS-based catalysts could also find utility in producing new indole derivatives; the indole ring is an extremely prevalent structure in pharmaceutical, agricultural and other industrial compounds.⁶ Recently, our laboratory revised the chemical mechanism for IGPS (Fig. 1).⁷ In the revised mechanism, Lys110 (numbering according to IGPS from the thermophilic organism *Sulfolobus sulfataricus*, ssIGPS) acts as a general acid in the ring closure step (encompassing condensation and decarboxylation) by donating a proton to the C2' carbonyl of CdRP. Dehydration is then facilitated through the general acid, Lys53, located on the $\beta 1\alpha 1$ loop, and the general base, Glu51, located on the end of the $\beta 1$ strand. The identification of Lys53 as the general acid in the dehydration step is intriguing because the $\beta 1\alpha 1$ loop is the most structurally dynamic loop in the enzyme according to molecular dynamic (MD) simulations.⁸ The conformational dynamics of this loop may also gate substrate binding to and product release from IGPS.^{9–11} As such, the $\beta 1\alpha 1$ loop plays a leading role in IGPS catalysis, and interactions with other regions of the enzyme, including other nearby loops, likely modulate the structure and/or dynamics of this loop and impact IGPS activity.

IGPS has the typical $(\beta/\alpha)_8$ -barrel, or triosephosphate isomerase (TIM)-barrel fold, which consists of eight parallel β -strands in the shape of a wheel that are surrounded by eight α -helices. Like other TIM-barrel proteins, critical amino acid residues for IGPS are located on the loops connecting β -strands to α -helices ($\beta\alpha$ loops) and on the β -strands themselves. Interactions between amino acid residues on the loops of other TIM-barrel enzymes have been implicated in controlling enzyme activity.^{12–15} Previously, Shen *et al.*⁸ identified potentially interacting residues on the $\beta 1\alpha 1$ and $\beta 2\alpha 2$ loops in IGPS using statistical coupling analysis (SCA) and MD simulations. SCA identifies residue positions that covary, and proposes that these residues are important for maintaining structure, allowing proper protein folding, and/or facilitating protein function.¹⁶ When used in conjunction with MD simulations, this method can identify functionally important amino acid pairs that may help to coordinate protein motions important for enzyme catalysis.¹⁷ A similar methodology also identified other covarying amino acids in IGPS, and mutagenesis studies confirmed that these residues are catalytically and/or structurally important.¹⁸ The covarying residues in the $\beta 1\alpha 1$ and $\beta 2\alpha 2$ loops may be especially important considering the many functional roles of the $\beta 1\alpha 1$ loop.

Table I. The Steady State Kinetic Parameters of WT and Variant ssIGPS Enzymes Indicate that the Interaction Between Arg54 and Asn90 Affects Multiple Steps of IGPS Catalysis

Variant	Temp (°C)	k_{cat} s ⁻¹	K_M (nM)	k_{cat}/K_M ($\times 10^6$ M ⁻¹ s ⁻¹)	$(k_{\text{cat}})_{\text{mut}}/$ $(k_{\text{cat}})_{\text{WT}}$	$(k_{\text{cat}}/K_M)_{\text{mut}}/$ $(k_{\text{cat}}/K_M)_{\text{WT}}$	$\Delta\Delta G^a$ (kcal/mol)	SVE	SDKIE
WT	25	0.16 ± 0.02	74 ± 38	2.2				1.0 ± 0.2	1.2 ± 0.2
	37	0.42 ± 0.04	88 ± 47	4.8				0.6 ± 0.3	5.8 ± 0.1
	75	0.67 ± 0.03	44 ± 9	15				-0.2 ± 0.1	3.6 ± 0.3
R54A	25	0.16 ± 0.01	145 ± 36	2.2	1.0	1.0	0.00	1.3 ± 0.4	n.d. ^b
	37	0.35 ± 0.04	95 ± 41	3.7	0.83	0.77	0.16	0.5 ± 0.1	5.0 ± 1.0
	75	1.2 ± 0.1	74 ± 13	16	1.8	1.1	-0.05	-0.2 ± 0.1	1.2 ± 0.2
N90A	25	0.04 ± 0.03	36 ± 21	1.0	0.25	0.45	0.47	0.7 ± 0.3	1.0 ± 0.3
	37	0.17 ± 0.02	74 ± 31	2.3	0.40	0.48	0.45	0.06 ± 0.1	5.2 ± 0.4
	75	0.55 ± 0.02	79 ± 10	7.0	0.82	0.47	0.52	-0.2 ± 0.1	1.6 ± 0.2
N90Q	37	0.30 ± 0.02	132 ± 31	2.3	0.71	0.48	0.45	0.5 ± 0.1	4.1 ± 0.8
	75	0.89 ± 0.06	117 ± 26	7.6	1.3	0.51	0.47	0.5 ± 0.1	1.7 ± 0.2
R54A/N90A	37	0.19 ± 0.06	23 ± 12	8.2	0.45	1.7	-0.33	0.0 ± 0.3	2.2 ± 0.3
	75	0.51 ± 0.08	107 ± 49	4.8	0.76	0.31	0.81	-0.1 ± 0.1	1.0 ± 0.2
F89A	75	0.06 ± 0.01	1200 ± 700	0.05	0.09	3.3 × 10 ⁻³	3.95	-0.23 ± 0.07	1.0 ± 0.2

^a $\Delta\Delta G = -RT \ln ((k_{\text{cat}}/K_M)_{\text{mut}}/(k_{\text{cat}}/K_M)_{\text{WT}})$.

^b n.d. is not determined.

Specifically, SCA-MD analysis suggested that interactions between Arg54 and Asn90 on the $\beta 1\alpha 1$ and $\beta 2\alpha 2$ loops, respectively, of the ssIGPS enzyme are important for enzyme function.¹⁹ The C_α to C_α distance between these residues is 6.1 Å, although the side-chains may make transient hydrogen bonds. Structural comparison of the substrate-bound and product-bound forms of ssIGPS, together with the structural modeling of the intermediate in the active site, also suggests that the anthranilate group of the substrate moves from one hydrophobic pocket to another during catalysis.^{7,20} It is important to note that the conserved Phe89 on the $\beta 2\alpha 2$ loop comprises part of the anthranilate binding pocket by making π - π interactions with the CdRP substrate. In addition to acting as a general acid, Lys53 is known to interact with both the anthranilate carboxyl group and the C3' hydroxyl.^{9,10,20} Therefore, the interaction between Arg54 and Asn90 may help to coordinate conformational changes in both the loops and the substrate or intermediate itself. Indeed, conformational transitions have been monitored by fluorophore-labeling of the $\beta 1\alpha 1$ loop and have indicated that the loop and/or the surrounding area undergo conformational changes on the same time-scale as enzyme catalysis.¹¹

In this article, we have examined the importance of the $\beta 2\alpha 2$ loop and its interaction with the $\beta 1\alpha 1$ loop in ssIGPS catalysis through the analysis of variants with amino acid substitutions at positions 54, 89, and 90. The thermostability of the ssIGPS enzyme allows kinetic and biophysical studies to be studied across a wider range of temperatures in order to further build relationships between protein flexibility and stability. Our results indicate that the $\beta 1\alpha 1$ - $\beta 2\alpha 2$ loop interactions through Arg54 and Asn90 are important for proper positioning and

function of the general acid, Lys53. The N90A substitution, in particular, results in substantial changes to the structural dynamics of the $\beta 1\alpha 1$ loop, as ascertained through solution-state nuclear magnetic resonance (NMR) studies. The dynamics of the $\beta 1\alpha 1$ loop are likely important for multiple stages in enzyme function, including aiding the transition of the substrate intermediate from one catalytic pocket to the other, positioning Lys53 for the dehydration reaction and facilitating the release of IGP product. These studies provide the structural, dynamic and functional rationale for the coevolution of residues on the $\beta 1\alpha 1$ and $\beta 2\alpha 2$ loops of IGPS enzymes.

Results and Discussion

Phe89 on the $\beta 2\alpha 2$ loop is important for substrate binding and chemical catalysis

Since we are interested in the role of active-site loops in IGPS catalysis, and previous studies resolved the role of the conserved residues on the $\beta 1\alpha 1$ loop and adjacent $\beta 1$ strand (i.e., roles for Lys53 and Glu51),⁷ we examined the role of the invariant Phe89 residue located on the $\beta 2\alpha 2$ loop. This residue makes a π - π interaction with the anthranilate moiety of CdRP. Therefore, it is not surprising that the F89A amino acid substitution led to an approximately 24-fold increase in the Michaelis constant, K_M (Table I). Perhaps more interesting is the eleven-fold decrease in the maximum turnover rate, k_{cat} , associated with this variant, which suggests that chemical catalysis may also be impacted by this residue change.

More insight into the binding or chemical steps reported on by k_{cat} can be derived from solvent viscosity effects (SVE) and solvent deuterium kinetic isotope effects (SDKIE). For example, SVEs are

present if substrate binding, product release, and/or a large viscosity-dependent conformational change in the enzyme are partially or fully rate-determining. In the case of ssIGPS, the SDKIE reports on the proton transfer step from Lys110 in the ring closure step.^{7,21} It should be noted that SVE and SDKIE analysis is required to tease apart the effects on the individual kinetic steps,^{7,21} especially considering that presteady state kinetics was unable to resolve individual micro-rate constants for the two chemistry steps.¹¹

For the F89A variant, the lack of a SVE suggests that a chemistry step is rate-determining (i.e., k_{cat} is reporting on a chemical step), and the lack of a SDKIE suggests that the dehydration step has now become fully rate-determining (i.e., the dehydration step is isotope-insensitive, see Ref. 7; Table I). In WT ssIGPS, the ring closure step is at least partially rate-determining, as indicated by the substantial SDKIE at 75°C.²¹ The findings with the F89A variant suggests that this amino acid substitution affects both substrate binding and catalytic events associated with the dehydration step. The change in the identity of the rate-determining step also suggests that the steady-state kinetic results underestimate the impact of the F89A substitution on the dehydration step.

Interaction between $\beta 1\alpha 1$ and $\beta 2\alpha 2$ loops through Arg54 and Asn90 is important for ssIGPS catalysis

Our studies with Lys53⁷ and Phe89 amino acid substitutions indicate that these $\beta 1\alpha 1$ and $\beta 2\alpha 2$ loop residues are important for ssIGPS catalysis. We were also interested in the interaction between these loops that may impact ssIGPS function. A previous combined SCA-MD analysis identified an important interaction between Arg54 and Asn90 that could potentially be involved in modulating the function of the $\beta 1\alpha 1$ and $\beta 2\alpha 2$ loops.⁸ To probe these interactions further, we made amino acid substitutions at both positions 54 and 90, and characterized the variant enzymes at multiple temperatures.

The R54A substitution did not result in any substantial change to the steady-state kinetic parameters at any of the temperatures assayed (Table I). At 75°C, R54A ssIGPS had a small increase to k_{cat} (~1.8 fold) compared to WT ssIGPS, indicating that it has a slightly different response to temperature compared to WT ssIGPS. In contrast, the N90A substitution led to a more substantial 5-fold decrease in k_{cat} at 25°C, but at higher temperatures (i.e., 75°C), the kinetic parameters for N90A ssIGPS were closer to WT ssIGPS. The temperature dependent activity differences likely reflect changes to the rate-determining step at the different temperatures, as we previously observed for WT ssIGPS.²¹

For WT ssIGPS, the rate-determining step at 25°C is product release as indicated by the SVE approaching the theoretical maximum of one, and consistent with previous suggestions.^{11,21,22} At higher temperatures (i.e., 75°C), there is no longer a significant SVE, but there is a substantial SDKIE, indicating that the ring closure step is at least partially rate determining. For the N90A variant, there was not a substantial SVE at 37°C in contrast to WT ssIGPS (Table I). This finding suggests that k_{cat} for the N90A variant likely reflects a chemical step, and the effect of the N90A substitution on chemistry is underestimated at the lower temperatures (i.e., for WT ssIGPS, $k_{\text{chemistry}} \gg k_{\text{cat}}$ at 37°C). At higher temperatures, there is not a substantial SVE for either N90A or WT ssIGPS (Table I), suggesting k_{cat} is reporting on a chemical event for both of these enzymes.

The interaction between Arg54 and Asn90 is involved in the proper function of the acid/base in the dehydration step of ssIGPS catalysis

To gain more insight into how amino acid substitutions on the $\beta 1\alpha 1$ and $\beta 2\alpha 2$ loops may impact enzyme catalysis, we examined SDKIEs and pH effects for the R54A and N90A variants. Both R54A and N90A ssIGPS have substantially reduced SDKIEs compared to WT enzyme at 75°C (Table I). These findings suggest that amino acid substitutions at these positions influence the chemical steps of enzyme catalysis, although they do not substantially change the overall catalytic rate at 75°C. The loss of the SDKIE is similar to what we have seen for previous amino acid substitutions (i.e., K53R and E51Q)⁷; akin to these variants, the R54A and N90A substitutions are likely decreasing the rate of the dehydration step, such that it becomes more rate-determining than the ring-closure step.

To further assess the role of Asn90 in the catalytic mechanism of ssIGPS, we also determined the kinetic parameters for the N90Q substitution. Glutamine was chosen because it may be a less deleterious substitution than alanine, and glutamine is found at this position in other IGPS enzymes, including IGPS from *Escherichia coli*. Since glutamine is longer by a single methylene group compared to asparagine, it is expected to make a less optimal interaction with the $\beta 1\alpha 1$ loop. Similar to the N90A variant, the steady-state kinetics for N90Q ssIGPS were substantially decreased compared to WT ssIGPS at lower temperatures, but k_{cat} of the N90Q variant approached that of WT ssIGPS at higher temperatures. The SDKIE on k_{cat} was also substantially reduced compared to WT ssIGPS, similar to what was found for the N90A variant.

The ascending and descending limbs of the pH rate profile in WT ssIGPS were previously attributed to the general base, Glu51, and the general acid, Lys53, respectively, in the dehydration step of the

Figure 2. Perturbation in the $\beta 1\alpha 1$ – $\beta 2\alpha 2$ loop interactions alters the pK_a of the general acid and base in the dehydration step of IGPS catalysis. pH rate profiles for (A) WT (pK_{a1} 5.6 ± 0.2 , pK_{a2} 8.7 ± 0.1), (B) R54A (pK_{a1} 6.5 ± 0.2 , pK_{a2} 8.5 ± 0.2), and (C) N90A (pK_{a1} 7.3 ± 0.3 , $pK_{a2} \gg 9$) ssIGPS enzymes show perturbations in the pK_a 's associated with acid/base catalysis in the dehydration step. This finding indicates that the N90A substitution is perturbing the dehydration step of the reaction, and interfering with proper function of the general acid Lys53.

catalytic mechanism.⁷ Therefore, the change in the pH rate profile induced by the N90A substitution (Fig. 2) suggests that this change disrupts the general acid activity of Lys53; this effect cannot be directly associated with Asn90 since this residue would not be a significant donor or acceptor of protons. It is likely that the interaction between the $\beta 1\alpha 1$ and $\beta 2\alpha 2$ loops helps to bind the substrate and/or helps to position Lys53 to act as the general acid. The proper structure of these loops and the surrounding area of the enzyme may also influence

the microenvironment of Lys53, which would also affect the pK_a of this residue.

A double mutant cycle is a classic way of determining if the function of two amino acid residues is coupled.²³ As such, we determined the effect of the R54A/N90A double substitution, and determined the free energy change induced on the catalytic efficiency compared to WT enzyme (i.e., $\Delta\Delta G = -RT \ln ((k_{cat}/K_M)_{mutant}/(k_{cat}/K_M)_{WT})$).²⁴ If residues at positions 54 and 90 are independent, then their thermodynamic effects should be additive (i.e., $\Delta\Delta G(R54A/N90A) = \Delta\Delta G(R54A) + \Delta\Delta G(N90A)$); nonadditive effects reveals coupling. Indeed, the effects of the R54A/N90A double substitution are not additive with the effects of the R54A and N90A single substitutions (Table I), suggesting some type of thermodynamic coupling between these residues. It should be noted that the N90A substitution is more deleterious than the R54A substitution. This finding might be explained by considering that the Asn90 side chain can hydrogen bond with the backbone of Arg54 and/or by noting that another residue on the $\beta 1\alpha 1$ loop (e.g., Lys55 or Ser56) may be free to make an interaction with the $\beta 2\alpha 2$ loop in the case of the R54A substitution. Altogether, these results indicate that the interaction between the $\beta 1\alpha 1$ and $\beta 2\alpha 2$ loops helps to increase catalytic efficiency by allowing residues that bind the substrate to make more favorable interactions, and by positioning the chemically relevant Lys53 to act as a general acid.

Structure and stability of loop variants

Covarying amino acids residues like Arg54 and Asn90 may be important not only for catalytic function, but may also be important for protein folding and stability.⁸ To assess the role of these amino acids for maintaining structure and thermal stability of ssIGPS, thermal inactivation experiments were performed for WT and variant ssIGPS enzymes. The thermal inactivation experiments showed a similar rate constant of inactivation for WT and R54A ssIGPS ($1.11 \times 10^{-3} \text{ s}^{-1}$ and $1.24 \times 10^{-3} \text{ s}^{-1}$, respectively) [Fig. 3(A)]. However, the two-fold decrease in the rate of inactivation observed for N90A ($6.91 \times 10^{-4} \text{ s}^{-1}$) and N90Q ($3.59 \times 10^{-4} \text{ s}^{-1}$) variants implies that changes to Asn90 increases the thermal stability of the enzyme. The circular dichroism (CD) spectra for WT and N90A ssIGPS enzymes [Fig. 3(B)] were nearly identical, indicating that the N90A substitution did not result in any gross structural change to the enzyme.

The structural dynamics of the $\beta 1\alpha 1$ loop are substantially altered in the N90A variant

Since the CD spectra did not indicate any gross structural change was induced by the N90A substitution, any kinetic changes associated with this amino acid substitution must be due to local structural and/or dynamic changes. Solution-state NMR

Figure 3. The N90A substitution affects the thermal stability but not proper folding for ssIGPS. (A) Thermal inactivation curves show an increase in the thermal stability of the N90A variant compared to WT ssIGPS. In brief, ssIGPS enzymes were incubated at 90°C for the times shown, and then assayed at 50°C at saturating CdRP concentrations. (B) Circular dichroism spectra indicate that the changes in activity for the N90A variant are not caused by gross structural changes to the enzyme and can be attributed to local changes to perturb the dehydration step.

offers powerful means to assess any local structure and dynamic changes in the $\beta 1\alpha 1$ and $\beta 2\alpha 2$ loops induced by the N90A substitution. As such, we assigned the backbone resonances for the $\beta 1$ strand, and $\beta 1\alpha 1$ and $\beta 2\alpha 2$ loops through standard triple-resonance NMR experiments (i.e., HNCA, HNCOCACB, HNCACB, HNCOCACB) and site-specific variants (i.e., A50P, Y52S, Y88S, and Y93S) for both WT and N90A ssIGPS enzymes (Supporting Information Figs. S1 and S2). Comparing the ^1H - ^{15}N heteronuclear single quantum coherence (HSQC) spectra for WT and N90A ssIGPS, we not only observe chemical shift changes around the site of mutation in the $\beta 2\alpha 2$ loop, but we also observe very substantial chemical shift changes for resonances associated with residues in the $\beta 1\alpha 1$ loop, especially for Lys53, Arg54, Lys55, and Ser56 [Fig. 4(A)]. There are only

relatively minor chemical shift changes for other resonances, suggesting that the structural/dynamic changes induced by the N90A variant are localized to these active-site loops.

To gain more insight into the structural/dynamic changes induced by the N90A substitution, we also determined the ^1H - ^{15}N heteronuclear Overhauser effects (hetNOEs) and R_{ex} , the contribution of conformational exchange on the μs -ms timescale to the R_2 relaxation rate, values. Experiments were conducted at 37°C, which coincides with the loss of the SVE for the N90A variant (Table I). In the absence of full resonance assignments, the hetNOE gives a qualitative estimate of motions on the ps-ns

Figure 4. Solution-state NMR studies indicate that the N90A substitution induces structural and dynamics changes to the $\beta 1\alpha 1$ loop. (A) Comparison of the ^1H - ^{15}N HSQC spectra for WT (black) and N90A (red) ssIGPS enzymes. Substantial chemical shift changes are observed for residues in both the $\beta 1\alpha 1$ (including residues 53–60) and $\beta 2\alpha 2$ (including residues 89–92) loops. (B) Comparison of the ^1H - ^{15}N heteronuclear Overhauser effects (hetNOE) for WT (black) and N90A (red) IGPS enzymes. Many of the residues in the $\beta 1\alpha 1$ loop have higher hetNOE values in the N90A variant compared to WT ssIGPS, suggesting that the N90A substitution induces structural ordering on the ps-ns timescale. Standard deviation in the hetNOE values were less than the size of the symbols used. NMR data were collected at 310 K on samples containing 1 mM protein in 25 mM potassium phosphate pH 7.0, 75 mM KCl, 1 mM EDTA, 1 mM DTT, 0.02% NaN_3 , and 10% (v/v) $^2\text{H}_2\text{O}$.

timescale and roughly correlates with the S^2 order parameter. Residues on the $\beta 1\alpha 1$ loop (especially residues 54–60) are associated with higher hetNOE values in the N90A variant compared to WT ssIGPS [Fig. 4(B)], suggesting that the N90A substitution induces more ordering of this loop on the ps-ns timescale. These effects are also observed on the longer timescale. In WT ssIGPS, Ser56, which is near the tip of the $\beta 1\alpha 1$ loop, experiences substantial conformational exchange on the μ s-ms timescale (i.e., $R_{ex} \sim 8 \text{ s}^{-1}$), but these exchange processes appear to be quenched for the N90A variant (i.e., $R_{ex} \sim 0 \text{ s}^{-1}$).

Conclusions

Our results provide an analysis of the interaction between the Arg54 and Asn90 residues on the $\beta 1\alpha 1$ and $\beta 2\alpha 2$ loops, respectively, of the tryptophan biosynthetic enzyme, ssIGPS (Fig. 1). This interaction was previously highlighted based on SCA-MD analysis.¹⁹ The SCA-MD method is proposed to identify residue pairs that may not be conserved but whose interactions still play integral roles in enzyme catalysis, especially in coordinating motions that may be important for enzyme function.^{16,25} Our kinetic analyses indicate that the Arg54-Asn90 interaction plays an important role in the catalytic mechanism of ssIGPS, particularly in the acid/base chemistry of the dehydration step. Although amino acid substitutions at these positions led to only modest changes in the overall steady-state kinetic parameters of ssIGPS, the resulting variants, in particular N90A ssIGPS, had a different rate-determining step than WT ssIGPS, especially at lower temperature (Table I). This result suggests that the changes in the steady-state kinetic parameters underestimate the impact of these amino acid substitutions on the individual rate constants. Perhaps most surprising were the changes induced by the N90A substitution on the pK_a s associated with the general acid and base of the dehydration step (Fig. 2). The N90A substitution did not result in any gross structural change to the ssIGPS enzyme (Fig. 3); instead our NMR studies suggest that the N90A substitution reduces structural dynamics of the $\beta 1\alpha 1$ loop across multiple timescales (Fig. 4). The NMR results are also consistent with the thermal inactivation studies, which suggest that the N90A variant is more thermally stable, and therefore likely less flexible, than WT ssIGPS (Fig. 3). $\beta 1\alpha 1$ loop dynamics may be controlled through competing noncovalent interactions; on the N-terminal side of the loop, Arg54 and potentially other residues make noncovalent interactions with Asn90, and on the C-terminal side of the loop, Arg64 and Asp65 make interactions with the residues in the $\alpha 1$ (i.e., Pro66, Ile67, Glu68, and Tyr69) and $\alpha 8'$ (i.e., Met237 and Arg238) helices. As such, there could be a type of “tug-of-war” between Asn90 and residues of the $\alpha 1/\alpha 8'$ helices that regulate the

structural dynamics of the $\beta 1\alpha 1$ loop. This type of effect has been referred to as protein “frustration.”²⁶ The loss of the interactions with Asn90 may remove this “frustration” and alter the structural dynamics of the $\beta 1\alpha 1$ loop. The conformational dynamics of the substrate and/or intermediates may also be adversely affected by changes in the $\beta 1\alpha 1$ and/or $\beta 2\alpha 2$ loops.

Changes to the conformational dynamics of the $\beta 1\alpha 1$ loop will also impact substrate binding/product release and the positioning and microenvironment of Glu51 and Lys53, which likely explains why product release is no longer rate-determining at 37°C and why the pK_a s of the general acid and base are substantially altered in the N90A variant. The N90A substitution is less deleterious at higher temperatures, probably because the thermal energy at higher temperatures allows IGPS to adequately sample the conformation(s) that is required for proper catalysis by Glu51 and Lys53. We have also suggested that there must be a conformational rearrangement in the intermediate for the dehydration reaction to occur⁷; Phe89 is likely an important residue guiding this rearrangement and the interaction between Arg54 and Asn90 likely coordinates the actions of Lys53 and Phe89. Altogether, we suggest that the interactions between the $\beta 1\alpha 1$ and $\beta 2\alpha 2$ loops govern multiple steps in the catalytic mechanism of ssIGPS, providing a functional rationale for the covariation of Arg54 and Asn90 in IGPS enzymes across multiple species.

Materials and Methods

Site-directed mutagenesis, overexpression, and purification of ssIGPS

Site-directed mutagenesis to achieve ssIGPS variants A50P, Y52S, R54A, Y88S, F89A, N90A, N90Q, and Y93S were performed using the QuikChange lightning site-directed mutagenesis kit (Agilent Technologies) and appropriate primers. All sequences (wild-type and mutant) were confirmed through DNA sequencing (Nucleic Acid Facility, Pennsylvania State University). Overexpression and purification of wild-type (WT) and variants ssIGPS enzymes was performed similarly as described,^{21,27} except proteins for NMR studies were isotopically labeled using established procedures (see below and Ref. 28).

Steady-state enzyme assays for WT and variant ssIGPS

Steady-state enzyme assays for ssIGPS followed previously established procedures.^{21,29} In short, ssIGPS activity was measured via fluorescence by monitoring the formation of IGP. IGP was excited at 278 nm and emission measured at 340 nm. Initial rate data

were fit to the Michaelis-Menten [Eq. (1)] using nonlinear regression with the program Kaleidograph:

$$v = (k_{\text{cat}}/E_T)/(K_M + [S]) \quad (1)$$

where v is the initial reaction velocity, E_T is the total amount of enzyme in the assay, and $[S]$ is the substrate concentration.

The solvent viscosity, solvent deuterium kinetic isotope and pH effects for WT and variant ssIGPS enzymes were determined by varying the buffer conditions of the standard assay as previously described.^{7,21} Solvent viscosity effects (SVE) were determined using enzyme assays in 50 mM HEPES, 4 mM EDTA, pH 7.5 buffer with glycerol (0 to 30% (w/v)). Studies with sucrose gave similar results as those with glycerol (data not shown). The relative viscosities of the buffer solutions were measured using an Ostwald viscometer. The addition of microviscogen did not change K_M values, but had effects on k_{cat} values where noted. PEG-8000 was used as a macroviscogen control, but did not alter the kinetic parameters of WT or variant enzymes (i.e., $k_{\text{cat,PEG-8000}}/k_{\text{cat,no viscogen}} = 1.0 \pm 0.2$ for 1.9% (w/v) PEG-8000 at a relative viscosity of 1.7). SVEs were obtained from the slope of a plot of relative viscosity versus $\text{rate}_0/\text{rate}_{\text{viscogen}}$.

Solvent deuterium kinetic isotope effects (SDKIE) were measured for k_{cat} in D_2O with WT and variant ssIGPS enzymes using saturating concentrations of CdRP (i.e., $[CdRP] > 5 \times K_M$). pD values were determined through measurement of the pH ($\text{pD} = \text{pH} + 0.4$). The SDKIE was defined as k_{H_2O}/k_{D_2O} .

pH studies for ssIGPS were performed using the following buffers: MES (pH 5.0–6.5), HEPES (pH 6.5–8.0), bicine (pH 8.0–8.5), and CHES (pH 8.6–10.0). None of the buffers gave any significant non-specific effects. The pH rate profiles were fit using Kaleidograph. For WT and R54A ssIGPS, which display two ionizations, data was fit to Eq. (2) through nonlinear regression:

$$v = C/(1 + 10^{pKa1 - pH} + 10^{pH - pKa2}) \quad (2)$$

For N90A ssIGPS, which displays only a single ionization, data was fit to Eq. (3) through nonlinear regression:

$$v = C/(1 + 10^{pKa1 - pH}) + D \quad (3)$$

where v is the estimated first-order (k_{cat}) or second-order (k_{cat}/K_M) rate constant, C is the maximum pH-independent rate value, D is the minimum pH-independent rate value, and pK_{a1} and pK_{a2} are the pK_a values associated with the ascending and descending limbs of the pH rate profile, respectively.

In our studies, pH changes did not substantially affect K_M values.

Thermal inactivation studies

ssIGPS enzymes (non His-tagged, WT and variant) were incubated at 90°C for up to 30 min. Aliquots were removed every 3 min and the enzyme activity was assayed in duplicate at 50°C with a saturating CdRP concentration (i.e., 800 nM). The data was fit to a linear regression of relative rate (compared to the rate prior to incubation) versus incubation time in seconds. The first-order rate constant of thermal inactivation, k_{inact} , is equal to the slope of this linear fit.

CD spectra of WT and variant IGPS enzymes

Circular dichroism (CD) spectra were recorded on a Jasco J-810 Spectropolarimeter from 250 to 190 nm with 1 nm intervals and a 1 nm bandwidth. The experiments were performed in 10 mM potassium phosphate pH 7.0 with an enzyme concentration of 1.7 μM .

NMR studies

For triple resonance NMR experiments, WT and N90A ssIGPS enzymes were overexpressed in M9 minimal medium containing $^{15}\text{NH}_4\text{Cl}$ and ^{13}C -glucose using standard procedures.²⁸ For relaxation experiments, WT and N90A ssIGPS were overexpressed in M9 minimal medium containing only $^{15}\text{NH}_4\text{Cl}$. The A50P, Y52S, Y88S, and Y93S variants, which were used to help assign the backbone resonances of ssIGPS, were overexpressed to achieve ^{15}N backbone-labeling of only Ala or Tyr residues.²⁸ For selective amino acid labeling, the appropriate ^{15}N -labeled amino acid was added to the overexpression media. In addition, 1 g/L of glyphosate was added 1 h prior to IPTG-induced protein expression ($\text{OD}_{600} \sim 0.4$ – 0.5) for ^{15}N Tyr labeling. Glyphosate inhibits aromatic amino acid biosynthesis and helps to prevent scrambling of the ^{15}N isotopic label. Isotopically-labeled proteins were purified according to previous procedures.²¹ After protein purification, ssIGPS enzymes were concentrated to 0.5–2 mM and buffer-exchanged into 25 mM potassium phosphate pH 7.0, 75 mM KCl, 1 mM EDTA, 1 mM dithiothreitol (DTT), 0.02% NaN_3 , and 10% (v/v) $^2\text{H}_2\text{O}$ using ZEBRA desalting spin columns (Thermo Scientific).

For backbone assignments, standard HNCA, HNCOCa, HNCACB and HNCOCACB spectra were obtained at 303 or 310 K on either a Bruker Avance III 600 MHz or a Bruker Avance III 850 MHz, both equipped with a TCI cryoprobe. All other NMR experiments were conducted at 310 K. ^1H - ^{15}N heteronuclear NOE values were measured by acquiring two spectra, with or without proton presaturation, in an interleaved manner. Two pairs of spectra were collected. For estimates of R_{ex} values, R_2 relaxation rates were measured on a Bruker Avance III 600

MHz spectrometer using constant-time relaxation-compensated CPMG pulse sequences as previously described.^{30,31} The effective R_2 relaxation rates (R_2^{eff}) were determined from the relation $R_2^{\text{eff}} = (-\ln(I(v_{\text{CPMG}})/I(0))/T$, where T ($=40$ ms) is the total relaxation period during CPMG pulsing, $I(v_{\text{CPMG}})$ represents peak intensities with CPMG pulsing at a particular radio frequency field, $v_{\text{CPMG}} = 1/\tau$, where τ is the time between successive 180° pulses in the CPMG sequence, and $I(0)$ is the intensity in the reference spectrum obtained without the CPMG pulse train. R_{ex} values were estimated as $R_{\text{ex}} = R_2^{\text{eff}}(v_{\text{CPMG}} = 100 \text{ s}^{-1}) - R_2^{\text{eff}}(v_{\text{CPMG}} = 2000 \text{ s}^{-1})$. $R_{2,0}$ is estimated as $R_2^{\text{eff}}(v_{\text{CPMG}} = 2000 \text{ s}^{-1})$ (i.e., R_2^{eff} at the highest 180° pulsing rate).

Electronic Supporting Information

Figures S1 and S2 showing NMR strip plots for the backbone resonance assignments for residues in the $\beta 1$ strand, $\beta 1\alpha 1$ loop and $\beta 2\alpha 2$ loop for the N90A variant (ZaccardiBoehr2013_supplementary.docx).

Acknowledgments

We thank Dr. Stephen Benkovic and Michelle Spiering for the kind use of their spectrofluorometer, and Melissa Mullen and Dr. Phillip Bevilacqua for their assistance with circular dichroism experiments.

References

- Shen H, Wang F, Zhang Y, Huang Q, Xu S, Hu H, Yue J, Wang H (2009) A novel inhibitor of indole-3-glycerol phosphate synthase with activity against multidrug-resistant *Mycobacterium tuberculosis*. *FEBS J* 276:144–154.
- Shen H, Yang E, Wang F, Jin R, Xu S, Huang Q, Wang H (2010) Altered protein expression patterns of *Mycobacterium tuberculosis* induced by ATB107. *J Microbiol* 48:337–346.
- Czekster CM, Neto BAD, Lapis AAM, Dupont J, Santos DS, Basso LA (2009) Steady-state kinetics of indole-3-glycerol phosphate synthase from *Mycobacterium tuberculosis*. *Arch Biochem Biophys* 486:19–26.
- Rothlisberger D, Khersonsky O, Wollacott AM, Jiang L, DeChancie J, Betker J, Gallaher JL, Althoff EA, Zanghellini A, Dym O, Albeck S, Houk KN, Tawfik DS, Baker D (2008) Kemp elimination catalysts by computational enzyme design. *Nature* 453:190–U194.
- Jiang L, Althoff EA, Clemente FR, Doyle L, Rothlisberger D, Zanghellini A, Gallaher JL, Betker JL, Tanaka F, Barbas CF, Hilvert D, Houk KN, Stoddard BL, Baker D (2008) De novo computational design of retro-aldol enzymes. *Science* 319:1387–1391.
- Barden TC (2011) Indoles: Industrial, agricultural and over-the-counter uses. *Top Heterocycl Chem* 26:31–46.
- Zaccardi MJ, Yezdimer EM, Boehr DD (2013) Functional identification of the general acid and base in the dehydration step of indole-3-glycerol phosphate synthase catalysis. *J Biol Chem* 288:26350–26356.
- Shen HB, Xu F, Hu HR, Wang FF, Wu Q, Huang Q, Wang HH (2008) Coevolving residues of (beta/alpha)(8)-barrel proteins play roles in stabilizing active site

- architecture and coordinating protein dynamics. *J Struct Biol* 164:281–292.
- Mazumder-Shivakumar D, Bruice TC (2004) Molecular dynamics studies of ground state and intermediate of the hyperthermophilic indole-3-glycerol phosphate synthase. *Proc Natl Acad Sci USA* 101:14379–14384.
- Mazumder-Shivakumar D, Kahn K, Bruice TC (2004) Computational study of the ground state of thermophilic indole glycerol phosphate synthase: structural alterations at the active site with temperature. *J Am Chem Soc* 126:5936–5937.
- Schlee S, Dietrich S, Kurcon T, Delaney P, Goodey NM, Sterner R (2012) Kinetic mechanism of indole-3-glycerol phosphate synthase. *Biochemistry* 52:132–142.
- Carpenter RA, Xiong J, Robbins JM, Ellis HR (2011) Functional role of a conserved arginine residue located on a mobile loop of alkanesulfonate monooxygenase. *Biochemistry* 50:6469–6477.
- Lipchick J, Loria JP (2009) Millisecond dynamics in the allosteric enzyme imidazole glycerol phosphate synthase (IGPS) from *Thermotoga maritima*. *J Biomol NMR* 45:73–84.
- Rozovsky S, Jogl G, Tong L, McDermott AE (2001) Solution-state NMR investigations of triosephosphate isomerase active site loop motion: ligand release in relation to active site loop dynamics. *J Mol Biol* 310:271–280.
- Rozovsky S, McDermott AE (2001) The time scale of the catalytic loop motion in triosephosphate isomerase. *J Mol Biol* 310:259–270.
- Lockless SW, Ranganathan R (1999) Evolutionarily conserved pathways of energetic connectivity in protein families. *Science* 286:295–299.
- Estabrook RA, Luo J, Purdy MM, Sharma V, Weakliem P, Bruice TC, Reich NO (2005) Statistical coevolution analysis and molecular dynamics: identification of amino acid pairs essential for catalysis. *Proc Natl Acad Sci USA* 102:994–999.
- Dietrich S, Borst N, Schlee S, Schneider D, Janda JO, Sterner R, Merkl R (2012) Experimental assessment of the importance of amino acid positions identified by an entropy-based correlation analysis of multiple-sequence alignments. *Biochemistry* 51:5633–5641.
- Shen H, Xu F, Hu H, Wang F, Wu Q, Huang Q, Wang H (2008) Coevolving residues of (beta/alpha)(8)-barrel proteins play roles in stabilizing active site architecture and coordinating protein dynamics. *J Struct Biol* 164:281–292.
- Hennig M, Darimont BD, Jansonius JN, Kirschner K (2002) The catalytic mechanism of indole-3-glycerol phosphate synthase: crystal structures of complexes of the enzyme from *Sulfolobus solfataricus* with substrate analogue, substrate, and product. *J Mol Biol* 319:757–766.
- Zaccardi MJ, Mannweiler O, Boehr DD (2012) Differences in the catalytic mechanisms of mesophilic and thermophilic indole-3-glycerol phosphate synthase enzymes at their adaptive temperatures. *Biochem Biophys Res Commun* 418:324–329.
- Merz A, Yee MC, Szadkowski H, Pappenberger G, Crameri A, Stemmer WPC, Yanofsky C, Kirschner K (2000) Improving the catalytic activity of a thermophilic enzyme at low temperatures. *Biochemistry* 39:880–889.
- Carter PJ, Winter G, Wilkinson AJ, Fersht AR (1984) The use of double mutants to detect structural changes in the active site of the tyrosyl-tRNA synthetase (*Bacillus stearothermophilus*). *Cell* 38:835–840.
- Ghosh A, Sakaguchi R, Liu C, Vishveshwara S, Hou YM (2011) Allosteric communication in cysteinyl tRNA

- synthetase: a network of direct and indirect readout. *J Biol Chem* 286:37721–37731.
25. Estabrook RA, Luo J, Purdy MM, Sharma V, Weakliem P, Bruice TC, Reich NO (2005) Statistical coevolution analysis and molecular dynamics: identification of amino acid pairs essential for catalysis. *Proc Natl Acad Sci USA* 102:994–999.
 26. Ferreira DU, Hegler JA, Komives EA, Wolynes PG (2007) Localizing frustration in native proteins and protein assemblies. *Proc Natl Acad Sci USA* 104:19819–19824.
 27. Hennig M, Darimont B, Sterner R, Kirschner K, Jansonius JN (1995) 2.0 Å structure of indole-3-glycerol phosphate synthase from the hyperthermophile *Sulfolobus solfataricus*: possible determinants of protein stability. *Structure* 3:1295–1306.
 28. Muchmore DC, McIntosh LP, Russell CB, Anderson DE, Dahlquist FW (1989) Expression and nitrogen-15 labeling of proteins for proton and nitrogen-15 nuclear magnetic resonance. *Methods Enzymol* 177:44–73.
 29. Schneider B, Knochel T, Darimont B, Hennig M, Dietrich S, Babinger K, Kirschner K, Sterner R (2005) Role of the N-terminal extension of the (betaalpha)8-barrel enzyme indole-3-glycerol phosphate synthase for its fold, stability, and catalytic activity. *Biochemistry* 44:16405–16412.
 30. Loria JP, Rance M, Palmer AG, 3rd (1999) A TROSY CPMG sequence for characterizing chemical exchange in large proteins. *J Biomol NMR* 15:151–155.
 31. Axe JM, Boehr DD (2013) Long-range interactions in the alpha subunit of tryptophan synthase help to coordinate ligand binding, catalysis, and substrate channeling. *J Mol Biol* 425:1527–1545.